Estimation of sampling and analytical uncertainties using Excel's 2-Way ANOVA with replications

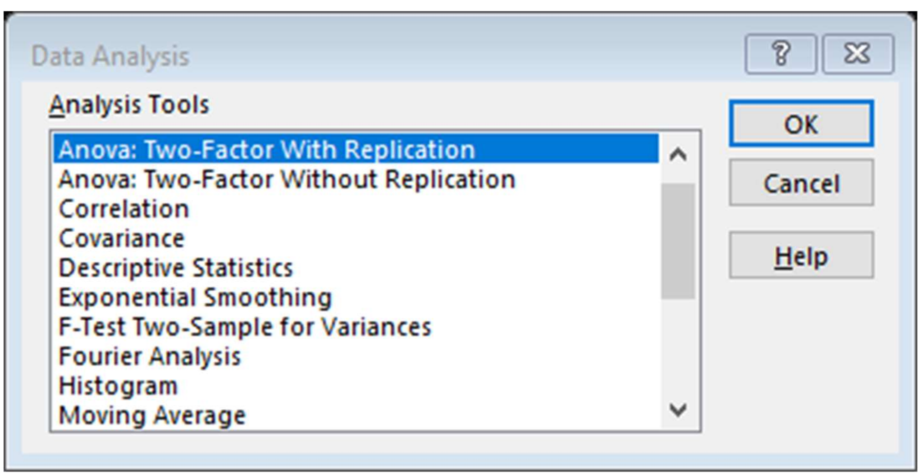
In the previous blog https://consultglp.com/2018/08/22/a-worked-example-of-measurement-uncertainty-for-a-non-homogeneous-population/ , , we used the basic ANOVA principles to analyze the total chromium Cr data for the estimation of measurement uncertainty covering both sampling and analytical uncertainty.

This article wants to show how you can make use of the Excel spreadsheet Data Analysis add-in to solve the similar issue, as it will save a good deal of your effort in various calculations.

As the previous study was about the duplicate analysis of duplicate samples from 10 random positions of a non-homogeneous polluted site, we shall choose "Anova Two-Factor with Replication" tool from the Excel spreadsheet list box when we select Data Analysis on the Ribbon's Data tab for our data evaluation.

It may be noted that this tool is intended for the sort of balanced two-factor design with *similar* number of replication of laboratory analysis. See Figure 1.

Figure 1: Selection of ANOVA method from the dialog box of the Excel Data Analysis Tool Package.



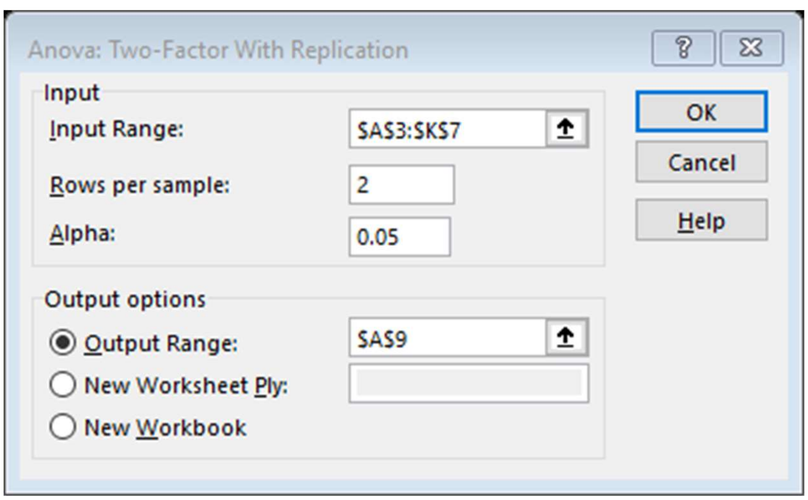
To suit the Excel's spreadsheet default settings, you need to present our data in the following format as shown in Table 1.

Table 1: Raw data of total chromium in mg kg^{-1} of samples in 10 locations

	A	B	C	D	E	F	G	H	I	J	K
1	**Excel Two-Factor ANOVA**										
2	**Total chromium Cr of duplicate samples from 10 different locations of a field**										
3		A	B	C	D	E	F	G	H	I	J
4	S1-A1	134	245	65	202	345	311	222	145	286	326
5	S1-A2	148	231	78	218	340	289	245	120	272	336
6	S2-A1	165	265	45	186	345	267	243	115	226	321
7	S2-A2	155	276	59	165	356	288	256	121	242	297
8											

Figure 2 shows the selection of data input in the Input Range cell with an entry of “2” keyed in the “Row per sample” cell as there are duplicated analysis done on each sample. After selecting the output range, the two-factor ANOVA outputs are shown in Figure 3.

Figure 2: A screen shot of Anova Two-Factor with Replication dialog box.
(Take note that some care has to be taken in the layout of raw data as shown in Figure 1 for use with this ANOVA tool.)



You might have noticed that there is no Labels in First Row check box in the dialog box shown in Figure 2. That is because you do not have a choice. The Input Range specified *must* include both row and column labels. That means, as the raw data set is laid out in Figure 1, you must specify the range : $A$3:$K$7 in the Input Range.

After selecting the locations where you want the output to start in the Output Range of Figure 2 and clicking “OK” to run the tool, the two-factor ANOVA results are a shown in Figure 3.

Figure 3: The outputs of Excel two-factor with replication on the total chromium Cr results from 10 locations of a field

Excel Two-Factor ANOVA
Total chromium Cr of duplicate samples from 10 different locations of a field

	A	B	C	D	E	F	G	H	I	J
S1-A1	134	245	65	202	345	311	222	145	286	326
S1-A2	148	231	78	218	340	289	245	120	272	336
S2-A1	165	265	45	186	345	267	243	115	226	321
S2-A2	155	276	59	165	356	288	256	121	242	297

Anova: Two-Factor With Replication

SUMMARY	A	B	C	D	E	F	G	H	I	J	Total
S1-A1											
Count	2	2	2	2	2	2	2	2	2	2	20
Sum	282	476	143	420	685	600	467	265	558	662	4558
Average	141	238	71.5	210	342.5	300	233.5	132.5	279	331	227.9
Variance	98	98	84.5	128	12.5	242	264.5	312.5	98	50	7772
S2-A1											
Count	2	2	2	2	2	2	2	2	2	2	20
Sum	320	541	104	351	701	555	499	236	468	618	4393
Average	160	270.5	52	175.5	350.5	277.5	249.5	118	234	309	219.65
Variance	50	60.5	98	220.5	60.5	220.5	84.5	18	128	288	8073.2
Total											
Count	4	4	4	4	4	4	4	4	4	4	
Sum	602	1017	247	771	1386	1155	966	501	1026	1280	
Average	150.5	254.25	61.75	192.75	346.5	288.75	241.5	125.25	256.5	320	
Variance	169.67	404.92	187.58	512.92	45.667	322.92	201.67	180.25	750.33	274	

ANOVA

ource of Variatic	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	680.63	1	680.63	5.2026	0.0337	4.3512
Columns	292589	9	32510	248.5	2E-18	2.3928
Interaction	5852.6	9	650.29	4.9707	0.0014	2.3928
Within	2616.5	20	130.83			
Total	301739	39				

It is noted that the Excel's outputs show exactly the same answers on "Columns (Locations)", "Within" and "Total" sources of variations, as those presented in the last blog based on the first ANOVA principles. By adding the figures of "Sample" and "Interaction" together, they return precisely the same results as the "Between-samples" calculation made in the last blog.

The MS(within samples) is the standard uncertainty in laboratory analysis. The approach in the evaluation of measurement uncertainty covering both sampling and analytical uncertainties that follows is exactly the same as shown in the previous article.

To conclude, you will certainly agree with me that this approach illustrated is much faster but it is still beneficial to learn the basic ANOVA principles first before indulging in this quicker method.